

# An exact, cache-localized algorithm for the sub-quadratic convolution of hypercubes

Oliver Serang  
 Freie Universität Berlin and  
 Leibniz-Institute of Freshwater Ecology and Inland Fisheries (IGB)  
 orserang@uw.edu

17 July, 2015

## Abstract

Fast multidimensional convolution can be performed naively in quadratic time and can often be performed more efficiently via the Fourier transform; however, when the dimensionality is large, these algorithms become more challenging. A method is proposed for performing exact hypercube convolution in sub-quadratic time. The method outperforms FFTPACK, called via `numpy`, and FFTW, called via `pyfftw` for hypercube convolution. Embeddings in hypercubes can be paired with sub-quadratic hypercube convolution method to construct sub-quadratic algorithms for variants of vector convolution.

## Introduction

Convolution between tensors  $x$  and  $y$  is defined as

$$(x * y)[k] = \sum_{i,j : k=i+j} x[i] \cdot y[j],$$

where  $i$ ,  $j$ , and  $k$  are integer vectors of the same dimension. Multidimensional convolution is important for computing the probability mass functions of sums and differences of joint probability distributions (or equivalently, sums and differences of multidimensional random variables) [1].

The naive algorithm for multidimensional convolution simply performs the Cartesian product between all  $x[i]$  and all  $y[j]$  (**Algorithm 1**). In practice, the naive method is rarely used and is frequently eschewed in favor of multidimensional fast Fourier transform (FFT) convolution. FFT convolution is a numerical method, which typically uses Cooley-Tukey method by padding each axis of  $x$  and  $y$  with zeros until it is a power of 2 and then further padding it with zeros until it reaches the next power of 2 [2]. Then the multidimensional FFT of the padded  $x$  and  $y$

---

**Algorithm 1 Naive multidimensional convolution**


---

```

1: procedure CONVOLVE( $x, y$ )
2:    $z \leftarrow \text{tensor}(x.\text{shape} + y.\text{shape} - 1)$  ▷ Initialize with zeros
3:   for  $i$  from  $(0, 0, \dots)$  to  $x.\text{shape}$  do
4:     for  $j$  from  $(0, 0, \dots)$  to  $y.\text{shape}$  do
5:        $z[i + j] += x[i] \cdot y[j]$ 
6:     end for
7:   end for
8:   return  $z$ 
9: end procedure

```

---

are performed (the multidimensional FFT is equivalent to performing the 1D FFTs of every row, column, *etc.* of the input tensor), multiplied element-wise, and then the multidimensional inverse FFT is performed (this is the conjugate of the multidimensional FFT of the conjugate of the input tensor). The convolution result will be found inside the result of the multidimensional inverse FFT [3].

For 1D vectors of length  $N \gg 1$ , the Cooley-Tukey FFT convolution is substantially faster, requiring  $O(N \log_2(N))$  steps, compared to the  $O(N^2)$  steps for the naive convolution based on the Cartesian product. For 2D matrices of shape  $(N_1, N_2)$ , the row FFTs will cost  $O(N_1 \cdot N_2 \log_2(N_2))$  and the column FFTs will cost  $O(N_2 \cdot N_1 \log_2(N_1))$  for an overall runtime in  $O(N_1 \cdot N_2 (\log_2(N_1) + \log_2(N_2))) = O(N_1 \cdot N_2 \log_2(N_1 \cdot N_2))$ ; therefore, FFT convolution still requires  $O(N \log_2(N))$  steps where  $N$  is the flat length of the 2D matrix; when  $N$  is large, this can be much faster than the  $O(N^4)$  steps required by naive convolution. However, for each problem listed here the dimensionality is a constant, and is therefore not included in the big-oh runtime. Furthermore, the column FFTs will not access contiguous blocks of memory, and therefore will not cache efficiently without special consideration.

But in the extreme case of convolving two  $D$ -dimensional hypercubes  $\{0, 1\}^D$  (*i.e.*, a tensor with  $D$  axes, which each have shape  $\{0, 1\}$ ), constants introduced by the dimensionality cannot be ignored: While the flat length  $N = 2^D$ , zero padding doubles each axis from 2 values to 4 values, and thus the zero padded tensor has flat length  $N = 4^D$  or equivalently,  $N = (2^2)^D = (2^D)^2$ . Thus the runtime of the Cooley-Tukey FFT convolution will be in  $O\left((2^D)^2 \log_2((2^D)^2)\right) = O\left((2^D)^2 \cdot D\right) = O(N^2 \log_2(N))$ . In comparison, the naive approach will require  $O\left((2^D)^2\right) = O(N^2)$  steps, and the runtime of convolving hypercubes via the Cooley-Tukey FFT can be resemble or exceed the runtime of a naive approach. Note that even though the multidimensional indices will be vectors of  $D$  booleans and require  $D$  machine steps for each step in the Cartesian product, these boolean indices can possibly be embedded into a single machine precision integer.

In terms of the flat length, both naive and Cooley-Tukey convolution of hypercubes will be in  $\Omega(N^2)$ . As an alternative to the multidimensional Cooley-Tukey

FFT, the hypercubes can be zero padded into a  $3^D$  tensor of the result size (rather than the  $4^D$  tensor used by the Cooley-Tukey method), and then the multidimensional FFT can be performed as several row, column, *etc.* 1D FFTs of length 3. These 1D FFTs cannot be performed via the Cooley-Tukey algorithm because their lengths are not powers of 2, but they can be performed as DFTs in  $3 \times 3$  steps. That algorithm will require  $O(3^D \log_2(3^D))$  steps.

This manuscript presents an alternative algorithm for the convolution of hypercubes. Unlike FFT convolution, this algorithm is exact and visits the indices row order, and is thus inherently more cache performant. The proposed method does not use FFT, but still achieves the same sub-quadratic asymptotic runtime as the non-Cooley-Tukey FFT of  $3^D$  hypercubes.

The proposed method resembles a hypercube variant of Karatsuba's method [4]; but where Karatsuba's method is for performing multiplication of integers of arbitrary length, this method is for convolving hypercubes of arbitrary dimension. Unlike Karatsuba's method, these hypercubes can also contain real or floating point values rather than integral values, because the recursion is applied to hypercubes of smaller dimension rather than Karatsuba's method of applying the recursion by splitting integral strings in half. Also in contrast to Karatsuba's fast integer multiplication method, the result space of two  $N$ -digit integers will have  $2 \cdot N - 1$  digits, whereas the result space of hypercube convolution in dimension  $D$  will be  $3^D$ , which is in  $\Omega(N^{1.585})$  (note that the exponent has been rounded). This means that an algorithm with the runtime of Karatsuba's method would be optimal.

The method can be used to simply and efficiently compute sums and differences of multiple high-dimensional probability mass functions where all axes are Bernoulli variables. Hypercube embeddings can be used to efficiently solve alternate forms of 1D convolution.

## Methods

Observe that a tensor convolution of dimension  $D$  can be performed by peeling off the first axis to produce convolutions of dimension  $D - 1$ :

$$\begin{aligned}
 z &= x * y \\
 \forall k_1, k_2, \dots, z[k_1, k_2, \dots] &= \sum_{i_1, i_2, \dots} x[i_1, i_2, \dots] \cdot y[k_1 - i_1, k_2 - i_2, \dots] \\
 &= \sum_{i_1} \sum_{i_2, \dots} x[i_1, i_2, \dots] \cdot y[k_1 - i_1, k_2 - i_2, \dots] \\
 &= \sum_{i_1} x[i_1] * y[k_1 - i_1],
 \end{aligned}$$

where tensors are stored in row-major format, so that  $x[i_1]$  returns a tensor of dimension  $D - 1$ . From this principle, hypercube convolution in dimension  $D$  can be solved by four hypercube convolutions of dimension  $D - 1$  as follows.

As shown by the general tensor convolution case above,  $k = i + j$  requires that the first index  $k_1 = i_1 + j_1$ . There are three cases for  $k_1$ :  $k_1 \in \{0, 1, 2\}$ .  $k_1 = 0$  requires  $i_1 = 0, j_1 = 0$ . Likewise,  $k_1 = 2$  requires  $i_1 = 1, j_1 = 1$ . Lastly,  $k_1 = 1$  can occur when either  $i_1 = 0, j_1 = 1$  or  $i_1 = 1, j_1 = 0$ . Thus,  $z = x * y$  can be solved by performing 4 hypercube convolutions of dimension  $D - 1$ :

$$\begin{aligned} z[0] &= x[0] * y[0] \\ z[1] &= (x[0] * y[1]) + (x[1] * y[0]) \\ z[2] &= x[1] * y[1], \end{aligned}$$

where each  $*$  operation performs a hypercube convolution of dimension  $D - 1$  and the base case simply returns the product between the two numeric values when  $D = 0$ . This leads to a runtime recurrence  $T(D) = 4 \cdot T(D - 1) + \Omega(2^D)$ , which has closed form  $T(D)$  in  $\Omega((2^D)^2)$ , leaving an algorithm that is still quadratic in  $N$ .

However, the fact that the tensors are hypercubes can be exploited: First compute  $z[0]$  and  $z[2]$  as above. Then compute the convolutions of marginals over the rows,  $t = (x[0] + x[1]) * (y[0] + y[1]) = x[0] * y[0] + x[0] * y[1] + x[1] * y[0] + x[1] * y[1]$ .  $t - z[0] - z[2] = x[0] * y[1] + x[1] * y[0] = z[1]$ . Thus, the  $D$  dimensional convolution can be computed in only 3 convolutions of dimension  $D - 1$ . This will be exact when the dynamic range of the values allows  $(a + b) - a = b$ .

The runtime of computing  $x[0] + x[1]$  is  $2^{D-1}$  (the same is true for  $y[0] + y[1]$ ) and the runtime of subtracting  $t - z[0]$  is  $3^{D-1}$  (the same is true when subsequently subtracting  $z[2]$ ); therefore, the runtime recurrence is defined:

$$\begin{aligned} T(D) &= 3 \cdot T(D - 1) + 2 \cdot 2^{D-1} + 2 \cdot 3^{D-1} \\ &< 3 \cdot T(D - 1) + 2 \cdot 3^D \\ &= 3 \cdot (3 \cdot T(D - 2) + 2 \cdot 3^{D-1}) + 2 \cdot 3^D \\ &\vdots \\ &= 3^D \cdot T(0) + 2 \cdot 3^D \cdot D \\ &= 3^D + 2 \cdot 3^D \cdot D \\ &< 3^{D+1} \cdot D. \end{aligned}$$

$N = 2^D$ , so the runtime will be bounded by  $3 \cdot D \cdot 3^{\log_2(N)} = 3 \cdot D \cdot N^{\log_2(3)} \leq 3 \cdot D \cdot N^{1.585} \in O(N^{1.585} \log_2(N))$ . All operations are performed by pairing two contiguous blocks of memory (much like a 1D in-place Cooley-Tukey FFT after bit-reversal has been performed), and so the proposed method is highly cache performant (**Algorithm 2**).

---

**Algorithm 2 Fast hypercube convolution in C++.** The algorithm is invoked `HypercubeConvolution<D>::apply(dest, x, y, flat_length_of_dest)`. The contents of `x` and `y` will be modified as the algorithm runs.

```
template <unsigned int D>
```

```

class HypercubeConvolution {
public:
    static void apply(double*__restrict__ const dest, double*
        __restrict__ const x, double*__restrict__ const y, const
        unsigned long three_to_d) {
        const unsigned long two_to_d_minus_one = 1ul<<(D-1);

        // Compute dest[0]:
        HypercubeConvolution<D-1>::apply(dest, x, y, three_to_d/3);

        // Compute dest[2]:
        HypercubeConvolution<D-1>::apply(dest + 2*three_to_d/3, x +
            two_to_d_minus_one, y + two_to_d_minus_one, three_to_d
            /3);

        // Compute x[0] + x[1] and y[0] + y[1]:
        unsigned int k;
        for (k=0; k<two_to_d_minus_one; ++k)
            x[k] += x[k + two_to_d_minus_one];

        for (k=0; k<two_to_d_minus_one; ++k)
            y[k] += y[k + two_to_d_minus_one];

        // Compute dest[1] = conv(x[0] + x[1], y[0] + y[1]) - dest
        // [0] - dest[2]
        HypercubeConvolution<D-1>::apply(dest + three_to_d/3, x, y,
            three_to_d/3);
        for (k=0; k<three_to_d/3; ++k)
            dest[k + three_to_d/3] -= dest[k];
        for (k=0; k<three_to_d/3; ++k)
            dest[k + three_to_d/3] -= dest[k + 2*three_to_d/3];
    }
};

template <>
class HypercubeConvolution<1u> {
public:
    static void apply(double*__restrict__ const dest, double*
        __restrict__ const x, double*__restrict__ const y,
        unsigned long) {
        dest[0] = x[0] * y[0];
        dest[1] = x[1] * y[0] + x[0] * y[1];
        dest[2] = x[1] * y[1];
    }
};

```

---

## Results

The runtime and accuracy of the divide and conquer method is compared to Python’s Cooley-Tukey-based `fftconvolve` routine from `scipy.signal` and to convolution via Python’s `numpy.fft.fftn` to tensors of the result shape  $3^D$ , which are both implemented in Fortran by `numpy`. The runtime and accuracy are also compared to FFTW on  $3^D$  tensors via the `pyfftw` package [5]. The proposed divide and conquer algorithm is implemented without complex numbers or any libraries using  $< 40$  lines of template recursive C++ code. The template recursive formulation is easily optimized, because the recursive calls can be unrolled and inlined by the compiler (it is compiled with `clang++-3.8 -Ofast`). For each  $D$ , three simulations were performed and the median runtime is reported (**Table 1**). FFTW is first run without timing it so that the expensive `plan` step (essentially a form of just-in-time compilation) used by FFTW is not included in the runtime.

## Discussion

The resulting divide and conquer algorithm produces the same results as with FFT convolution, but with exact results and with a significantly faster runtime and lower memory footprint. Furthermore, the increased precision of the proposed algorithm means it can be paired with  $p$ -norm rings to more accurately approximate fast algorithms on the semiring  $(\times, \max)$  (*e.g.*, max-convolution), because much greater values of  $p$  will be numerically stable [6]. This means that max-convolution on hypercubes containing integers of bounded dynamic range can be performed exactly by using a value of  $p$  large enough that the relative error  $1 - N^{-\frac{1}{p}}$  drops low enough that the absolute error on the range of integers is  $< 0.5$ , and so rounding  $p$ -norm estimates to the nearest integer will be exact [7].

Hypercube embeddings can be used to perform variants of 1D vector convolution. For example, the proposed hypercube convolution method solves the 1D “carry-free convolution”. Where 1D convolution defines  $z[k] = \sum_{i,j : k=i+j} x[i] \cdot y[j]$ , the carry-free convolution would describe the same problem where the additions between integers apply no bitwise carry operations. For instance,  $i = 7$  and  $j = 5$  would add bitwise  $(1, 1, 1) + (1, 0, 1)$  to produce  $(2, 1, 2)$ , which would produce  $(1, 1, 0, 0)$  after carry operations and is equivalent to  $8 + 4 = 12 = 7 + 5$ . In the carry-free variant, the carry operations would not be performed during convolution; therefore,  $x[(1, 1, 1)] \cdot y[(1, 0, 1)]$  would be added to  $z[(2, 1, 2)]$  rather than to  $z[(1, 1, 0, 0)]$ . This seemingly innocuous change to 1D convolution means that it is not trivial to solve it efficiently using existing 1D convolution algorithms, but it can be solved by embedding into a hypercube of dimension  $D = \log_2(N)$ , where each axis corresponds to a  $\{0, 1\}$  value of the bits for a given index, and then convolving the two hypercubes. That hypercube convolution can be efficiently and accurately solved by the method proposed. It is likely that there are similar embeddings for other

$D$	11	12	13	14	15	16	17	18
$N$	2048	4096	8192	16384	32768	65536	131072	262144
	Runtimes (seconds)							
<code>fftconvolve</code>	0.930	4.00	17.1	81.0	—	—	—	—
<code>fftn</code> of $3^D$	0.078	0.266	0.902	3.19	10.2	35.5	115	—
FFTW of $3^D$	0.0136	0.0358	0.181	0.420	1.89	4.19	13.4	—
D&C	<b>0.00447</b>	<b>0.0155</b>	<b>0.0396</b>	<b>0.0429</b>	<b>0.167</b>	<b>0.444</b>	<b>1.39</b>	<b>4.56</b>
	Relative error at smallest element							
<code>fftconvolve</code>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	—	—	—	—
<code>fftn</code> of $3^D$	$1 \cdot 10^{-8}$	$1.7 \cdot 10^{-7}$	$1.08 \cdot 10^{-6}$	$1.52 \cdot 10^{-5}$	$1.23 \cdot 10^{-4}$	$3.95 \cdot 10^{-4}$	$2.98 \cdot 10^{-3}$	—
FFTW of $3^D$	$4 \cdot 10^{-8}$	$1.8 \cdot 10^{-7}$	$1.32 \cdot 10^{-6}$	$1.52 \cdot 10^{-5}$	$1.76 \cdot 10^{-5}$	$6.62 \cdot 10^{-4}$	$8.35 \cdot 10^{-3}$	—
D&C	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>	<b>0</b>

Table 1: **Runtimes and accuracies on hypercube convolutions of dimension  $D$ .** Runtimes (in seconds) computed using the `time.time` routine in Python and `std::clock_t` time in C++. Note that `fftconvolve` and `fftn` both call FFTPACK in Fortran, and that `fftconvolve` is able to easily exploit the fact that the convolution is on the reals because it operates on the  $4 \times 4 \times \dots$  tensor (whose axes are powers of 2). The `fftn`  $3^D$  method only zero pads into a  $3 \times 3 \times \dots$  tensor. FFTW uses multiple cores. All FFT methods perform row, column, *etc.* DFTs to perform the forward and inverse tensor FFTs and numerically compute the convolution. Values are unreported when more than the 16GB RAM available was required. Accuracy is evaluated using hypercubes with flat vectors  $x.flat = y.flat = [1, 2, 3, \dots, 2^D]$  and analyzing the relative error at the smallest result value  $(x * y)[(0, 0, \dots, 0)] = 1$ , which can suffer the greatest influence from large values elements in the convolution. Note that because `fftconvolve` is run on  $4 \times 4 \times \dots$  tensors, the twiddle factors are computed using angles of the form  $k \cdot \frac{\pi}{2}$ ,  $k \in \mathbb{Z}$  and thus accumulate no error, whereas applying FFTs to  $3 \times 3 \times \dots$  tensors yields larger errors from angles of the form  $k \cdot \frac{\pi}{3}$ . Even though it uses only one core, the proposed divide and conquer algorithm (labeled D&C) simultaneously achieves the highest accuracy and the fastest runtime (best results in bold font).

combinatoric problems that would benefit.

## References

- [1] O. Serang. The probabilistic convolution tree: Efficient exact Bayesian inference for faster LC-MS/MS protein inference. *PloS one*, 9(3):e91507, 2014.
- [2] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Mathematics of computation*, 19(90):297–301, 1965.
- [3] J. G. Proakis and D. G. Manolakis. *Introduction to digital signal processing*. Prentice Hall Professional Technical Reference, 1988.

- [4] A. Karatsuba and Y. Ofman. Multiplication of multidigit numbers on automata. In *Soviet physics doklady*, volume 7, page 595, 1963.
- [5] Matteo Frigo. A fast Fourier transform compiler. In *ACM SIGPLAN notices*, volume 34, pages 169–180. ACM, 1999.
- [6] O. Serang. A fast numerical method for max-convolution and the application to efficient max-product inference in Bayesian networks. *Journal of Computational Biology*, 22:770–783, 2015.
- [7] J. Pfeuffer and O. Serang. A bounded  $p$ -norm approximation of max-convolution for sub-quadratic Bayesian inference on additive factors. *Journal of Machine Learning Research*, 17(36):1–39, 2016.